COMPARING THE EFFECTS OF NARRATIVE NONFICTION AND LITERARY FICTION ON EMPATHY RETENTION IN MEDICAL STUDENTS

A thesis submitted to the University of Arizona College of Medicine – Phoenix
in partial fulfillment of the requirements for the Degree of Doctor of Medicine

Aishan Shi
Class of 2019

Mentor: Jennifer Hartmark-Hill, MD

ABSTRACT

Integration of medical humanities into medical student curricula has been shown to improve medical student empathy and resilience. The purpose of this study is to determine if narrative nonfiction pieces help students retain equal or more empathy skills compared to reading literary fiction. Previous studies show that interventions that utilize medical humanities can vary in medium and genre and face the challenge of small sample size and confirmation bias due to a lack of randomized trials. In contrast, this study compares the reading of Narrative Nonfiction and Literary Fiction in building empathy in second year medical students randomized to each genre. Participants were asked to read selections from their assigned genre during the intervention period. Baseline, pre-intervention, and post-intervention assessments were measured by the Reading the Mind in the Eyes – Revised. Wilcoxon Rank Sum was performed to compare continuous variables. Additionally, a linear mixed model was utilized to determine mean differences in scores over time for the Overall class, Fiction, and Nonfiction subgroups. Results demonstrated a statistically significant decrease in empathy across the overall study period, and there was no empathy retention difference between genres. Female gender identity and increased engagement in the arts and humanities prior to medical school were correlated with higher empathy scores across time. These findings indicate the need for longitudinal and personalized learning in medical humanities for more thorough studies and maximized benefits on empathy retention.

Table of Contents:

List of Figures and Tables:

INTRODUCTION, SIGNIFICANCE, AND RATIONALE

Building empathy as a means of increasing resilience and reducing burnout

Physician burnout is a growing concern, and early interventions to reduce burnout have become increasingly recognized as critical to trainee development.[1] It has been suggested that curricular activities that focus on empathy building can increase medical student and physician resilience (and therefore decrease burnout) by reinforcing the patient-physician relationship. Since the 1990s, medical educators have called for the inclusion of medical humanities into undergraduate medical curricula.[2] However, only a handful of studies have been performed to assess the outcomes of exposing medical students to patient or physician narrative literature and other literary genres.[6]

Mechanisms for building empathy in medical students

Interventions aimed at building empathy in medical students that utilize medical humanities can vary in medium and genre. Studies have used poetry,[3, 4] prose -- in the form of fiction,[4-6] patient or physician narratives (nonfiction) [5, 6] -- discussions,[5, 6] and blogging and/or reflective writing.[6] Other recently developed programs have also integrated art, dance, film, and other forms of more abstract expression into their studies.[6] These interventions have received positive responses from medical students, increased self-reported empathy and resilience, and demonstrated higher empathy assessment scores.[3-7]

The ability of patient narratives to build empathy is accomplished through a variety of factors. In a 2005 article, Alan Bleakly proposed that the process of thinking about narrative, thinking with narrative, and finally thinking through narrative take a reader from analyzing the sequence of the account as data (i.e. putting together a history to arrive at a differential) to understanding the emotional burden and the person in the patient.[8] In the first process (thinking about narrative), narrative has the ability to make the unfamiliar (a sequences of signs and symptoms) familiar (a diagnosis). However, in the last process (thinking through narrative), it takes the categorically familiar (disease in a patient) and turns it unfamiliar (a person). This

takes the analytical process one step further to the more complex phenomenology of consciousness and being.

Subsequently, in 2013, Bleakly and Marshall assert that humanities promote creative thinking and comfort with the abstract while simultaneously requiring clear expression.[9] As a result, students of the humanities have a greater tolerance for ambiguity and the unfamiliar, but also know how to navigate ambiguity with the appropriate exploratory dialogue, thereby improving communication. Ultimately, it is this dual nature of narrative to be able to provide many levels of 'truth' that makes a useful tool in developing empathy because it forces perspective-taking by the reader in order to simultaneously develop and accept these many levels of 'truth,' especially in the realm of diverse patient perspectives with regard to experiences and values for their care.[10]

Large-scale and long-term medical humanities studies

Chen and Forbes provide a useful systematic review of the different studies that have already completed, and the different assessment tools that have been used.[11] Nonetheless, measuring empathy and the outcomes of integrating medical humanities into curricula is not standardized. It is difficult for these assessments to accurately reflect empathy development. However, oftentimes, positive outcomes have resulted from medical humanities studies.[12] Other challenges that this type of research faces include confirmation and selection bias and small sample size because these are not performed as randomized trials; therefore, students with more interest in the humanities (and perhaps possess more humanistic characteristics) are more likely to be the ones who participate in these studies. Thus, one goal of this project is to study the effect of medical humanities exposure on a large sample size (n=80, an entire class of medical students) to assess the differences in empathy level.

Comparing narrative nonfiction and literary fiction

There is a need to assess the types of literature that is being presented and taught in a medical humanities curriculum and determine the mechanism of their effectiveness. Kidd and Castano published an article in 2013 asserting that reading literary fiction (which focuses on character

and challenges readers to learn about the “circumstances and inner lives of people unlike themselves”) can improve a person’s ability to empathize.[13] They cite Barthe's argument that literary fiction is a “writerly” genre rather than a “readerly” genre. Hence, literary fiction brings readers into the mind of the writer. The authors also reference literary theorist Bahktin who considers literary fiction to be a “cacophony of voices … [with] an absence of a single authorial perspective.” The selection of works used in Kidd and Castano’s study, as well as their use of both Barthe and Bahktin’s theories, leave many questions unanswered. Although the critical theories support the ability of literary fiction to evoke empathy from readers, it does not entirely rule out nonfiction as being unable to have this same effect. In particular, narrative nonfiction, such as memoirs and personal essays, often has the same features as literary fiction pieces.

This genre of Nonfiction was unexamined in Kidd and Castano’s study, leaving significant opportunity and impetus for further exploration. However, because medical humanities often fall into this category, it is important to compare the effect of narrative nonfiction to literary fiction. Furthermore, because narrative nonfiction in particular is not only a growing genre in literature, but it also a predominant feature of popular media such as blogs, radio and podcasts, magazines, etc., it may be more accessible and connectable for medical students than traditional literary pieces. It is also important to determine if short story forms of literary fiction produce the same outcomes as the novels that Kidd and Castano used. Shorter pieces are not only more digestible for students, but also serve as a useful validation tool for the mechanisms that allow literary fiction to build empathy in readers because length should not be a factor.

Research Methods

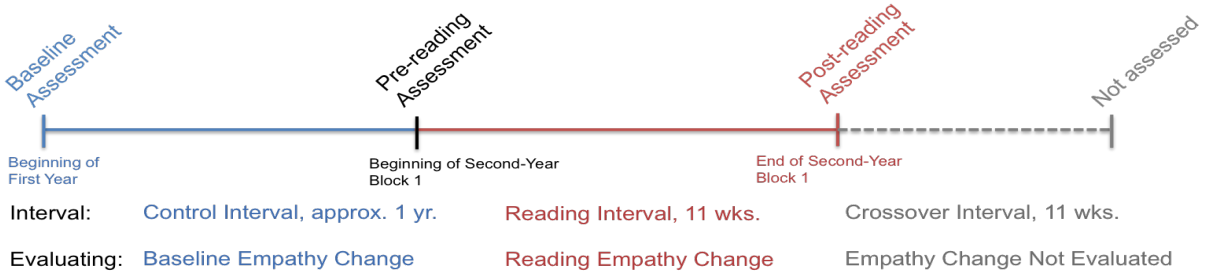
The aim of this study is to determine if reading Narrative Nonfiction pieces pertaining to medical experiences help students develop or retain equal or more empathy skills compared to reading Literary Fiction. IRB exemption was obtained, and students were able to anonymously provide or withhold consent for de-identified, aggregated assignment data to be used for further research analysis.

Subjects and Timeline

Medical students at the University of Arizona College of Medicine – Phoenix took a preliminary empathy assessment within the Capstones Course curriculum at the beginning of the first year. This served as a baseline at matriculation into medical school. After a one-year interval without designated reading assignments, a second assessment was collected at the beginning of second year to establish a Baseline Empathy Change (Pre-reading score minus Baseline score) that reflects expected fluctuations due to both stressors and positive influences on empathy related to medical school and other lived experiences. The reading assignment was then given out with students randomized into each genre. The Reading Interval spanned over the course of their first second-year basic sciences block (eleven weeks), and students were asked to read selections from their assigned genre. At the end of the Reading Interval, students took a final assessment of their empathy levels, and their Reading Empathy Change was calculated (Post-reading score minus the Pre-reading score).

The entire time frame of this study is considered the Study Period, and the score difference between the Post-reading and Baseline time points reflects the overall change in empathy score during this period. Following the reading intervention, students were then crossed-over into the alternative genre group for another eleven weeks to maintain educational equality; however, a second Post-reading empathy level was not assessed due to the likelihood of survey fatigue affecting the results. These time points, intervals, and evaluated empathy changes are depicted in Figure 1.

Figure 1: Timeline of the Study



The solid line represents the entire Study Period, and dashed line represents the crossover.

Selected Readings

Pieces from both the Narrative Nonfiction and Fiction genres were selected from medical and non-medical anthologies, with input from content experts in the Office of Diversity & Inclusion. Although the aim was to have equal representation from both medically related and unrelated literature, all of the texts could be read through the lens of medicine with regard to diverse patient populations. Selections ranged in length, but all participants were asked to read a minimum of three pieces for adequate exposure to the genre. The titles, authors, and page numbers from each genre are presented in Tables 1 and 2.

Table 1: Narrative Nonfiction Selections		
Title	Author	Page Numbers
An Anthropologist on Mars (excerpt)	Oliver Sacks	23
Black Man in a White Coat	Damon Tweedy	5
Doctor, Talk to Me	Anatole Broyard	10
Drinking: A Love Story	Caroline Knap	19
The Story of My Life	Helen Keller	19
What Doctors Feels	Danielle Ofri	22

Table 2: Literary Fiction Selections		
Title	Author	Page Numbers
All Boy	Lori Ostlund	17
Bless Me, Última (excerpt)	Rudolpho Anaya	23
Into Silence	Marlin Barton	22
Ordinary People	Judith Guest	26
Someone Ought to Tell Her There's Nowhere to Go	Danielle Evans	18
The Bonesetter's Daughter	Amy Tan	9

Assessment Tool: Reading the Mind in the Eyes Test - Revised (MIE)

The ability to interpret others' emotions is necessary for providing the appropriate empathetic response. This skill is formally known as Theory of Mind and is often referred to as social or emotional intelligence. Baron-Cohen, *et al.*, developed a revised Reading the Mind in the Eyes Test that utilizes 36 images of facial expressions focused on the eyes and requires subjects to match the emotion expressed with one of four multiple choice words. Foil words have similar emotional valence as the target words in order to test the ability of the subject to make these subtle distinctions. This test has been shown to be negatively correlated to the Autism Quotient. Average scores for general population are 26-28, with scores above 30 indicating higher emotional intelligence and empathetic recognition. This tool can be accessed in the Supplemental Data as Supplemental Table 1.[17]

Debriefing Session

Two debriefing sessions were held separately with each reading group after the participants took the Post-reading assessment. The purpose of the debrief sessions was to hold a discussion with students and help them further understand how their readings could help them reflect on their experiences and build empathy during and beyond medical school. These sessions were facilitated with the team from the Office of Diversity and Inclusion (ODI) as well as faculty and staff in the Department of Bioethics & Medical Humanism. This collaboration with ODI was helpful in teaching students bias awareness as a tool for perspective taking (a component of empathy). Students were asked to identify biases from the perspectives of various characters, narrators, and themselves as readers of the texts, and then consider how these biases affect their perceptions of the characters and situations in the readings. Because the Post-reading assessment was conducted prior to these debrief sessions, these sessions have no impact on the outcome of the reading interventions. Although no quantitative data were collected from this discussion activity, student feedback for the sessions was collected in an anonymous fashion as part of the Capstones Course evaluation.

Statistical Analysis

Statistical analysis was performed using STATA. Wilcoxon Rank Sum was performed to compare continuous variables. Linear regression was used to ascertain mean differences in the change in scores between time points adjusting for all other variables in the model. Additionally, a linear mixed model was utilized to determine mean differences in scores over time for the Overall class, Fiction, and Nonfiction subgroups.

Results

From a class of 80 students, 65 were successfully tracked through the completion of Baseline, Pre-reading, and Post-reading assessments, for a completion rate of 81.3%. Thirty-three students were randomized into the Nonfiction genre and 32 into the Fiction genre. Demographic information taken at Baseline asked participants to self-identify age range, gender, field of undergraduate study, and their level of engagement with arts and humanities before, while applying to, and during medical school. The breakdown of these demographics and the randomization of participants into the Fiction and Nonfiction genres are presented in Table 3. No statistically significant differences in the two groups were found across demographic variables.

Table 3: Demographics

	Overall N=65	Nonfiction N=33	Fiction N=32	P-value
Age, years (n ,%)				
≤25	45 (69.2)	26 (78.9)	19 (59.4)	0.09
>25	20 (30.8)	7 (21.2)	13 (40.6)	
Gender (n ,%)				
Female	39 (60.0)	23 (69.7)	16 (50.0)	0.11
Male	26 (40.0)	10 (30.3)	16 (50.0)	
Undergrad Multiple Fields (n ,%)				
No	47 (72.3)	23 (69.7)	24 (75.0)	0.63
Yes	18 (27.7)	10 (30.3)	8 (25.0)	
Undergrad Major (n ,%)				
STEM	43 (66.2)	22 (66.7)	24 (65.6)	0.43
Social Sciences	9 (13.9)	3 (9.09)	6 (18.8)	
Both	13 (20.0)	8 (24.2)	5 (15.6)	
Engaged in Arts and Humanities Before Medical School (n ,%)				
Never	3 (4.62)	2 (6.06)	1 (3.13)	0.17
Rarely	11 (16.9)	4 (12.1)	7 (21.9)	
Sometimes	24 (36.9)	10 (30.3)	14 (43.8)	
Often	19 (29.2)	12 (36.4)	7 (21.9)	
Very Often	8 (12.3)	5 (15.2)	3 (9.38)	
Engaged in Arts and Humanities while Applying to Medical School (n ,%)				
Never	3 (4.62)	2 (6.06)	1 (3.13)	0.051
Rarely	14 (21.5)	3 (9.09)	11 (34.4)	
Sometimes	25 (38.5)	12 (36.4)	13 (40.6)	
Often	18 (27.7)	15 (45.5)	3 (9.38)	
Very Often	5 (7.69)	1 (3.03)	4 (12.5)	
Engaged in Arts and Humanities While in Medical School (n ,%)				
Never	5 (7.69)	2 (6.06)	3 (9.38)	0.82
Rarely	29 (44.6)	14 (42.4)	15 (46.9)	
Sometimes	24 (36.9)	15 (45.5)	9 (28.1)	
Often	5 (7.69)	2 (6.06)	3 (9.38)	
Very Often	2 (3.08)		2 (6.25)	

Wilcoxon Rank Sum to compare ordinal variables. Chi-squared/Fisher's Exact to compare categorical variables.

Mean empathy scores (standard deviation) at Baseline for the Overall class, Nonfiction, and Fiction groups were 26.7 (4.57), 26.1 (5.56), and 27.4 (3.25), respectively, demonstrating no significant difference at this time point between the two genre groups (p-value 0.58). These scores are also consistent with the mean scores of the general population, as described by Baron-Cohen, *et al.*[17] At the Pre-reading time point one year later, empathy raw scores decreased for the Overall class and Fiction groups, and generally remained the same for the Nonfiction group; this finding was not of significance (p-value 0.91) as demonstrated in Table 4. The Baseline Empathy Change between the first and second year of medical school (Table 4) was also not significant between the two genres (p-value 0.16).

Following the eleven-week reading intervention, the Post-reading empathy scores decreased again (mean empathy raw scores [SD] of 24.6 [5.60] Overall, 24.2 [5.79] Nonfiction, and 25.0 [5.45] Fiction, and p-value 0.46 for scores between genre groups). The Study Period Empathy Changes (mean [SD]) were -1.95 [5.69], and -2.36 [4.69] in the respective Nonfiction, and Fiction groups, and Overall of -2.15 [5.19] points. Between genres, Study Period Empathy Change was also found to be not significant (p-value 0.56). However, this difference between baseline empathy score and post-reading empathy score was found to be significant over time for both the Overall class (p-value < 0.001) and in those randomized to each group (p-value 0.021 for Nonfiction and 0.001 for Fiction) as demonstrated at the bottom of Table 6.

Table 4: Empathy Scores and Empathy Changes Across Time Points

	Overall N=65	Nonfiction N=33	Fiction N=32	P-value
Baseline Raw Score (mean, SD)	26.7 (4.57)	26.1 (5.56)	27.4 (3.25)	0.58
Pre-Read Raw Score (mean, SD)	25.9 (4.57)	26.2 (4.04)	25.6 (5.10)	0.91
Post-Read Raw Score (mean, SD)	24.6 (5.60)	24.2 (5.79)	25.0 (5.45)	0.46
Baseline Empathy Change (mean, SD)	-0.83 (4.54)	0.08 (4.79)	-1.76 (4.15)	0.16
Reading Empathy Change (mean, SD)	-1.32 (3.84)	-2.03 (4.15)	-0.59 (3.40)	0.22
Study Period Empathy Change (mean, SD)	-2.15 (5.19)	-1.95 (5.69)	-2.36 (4.69)	0.56

Wilcoxon Rank Sum to compare continuous variables. The Baseline Empathy Change represents the difference between the Pre-reading time point and Baseline time point, whereas the Reading Empathy Change represents the difference between the Post-reading time point and Pre-reading time point. The Study Period Empathy Change represents the difference between the Post-read time point and Baseline time point.

Demographic variables and level of engagement with arts and humanities generally had no correlation with Baseline Empathy Change or Reading Empathy Change, with the exception of the significant finding that self-identification as Male correlated with a more negative Baseline Empathy Change compared to self-identification as female (-2.93 [CI -5.39 to -0.46], p-value 0.02). Male participants were also found to have consistently lower scores across all three time points, with an average of -2.75 points (CI -4.92, -0.59) and p-value 0.012, as seen in Table 6.

With regards variables affecting empathy changes over time, increased engagement in the arts and humanities prior to medical school consistently correlated with higher scores across time points. Participants who marked their engagement in arts and humanities prior to medical school as “Often” scored a mean difference of 9.21 (CI 0.23, 18.2) for the Overall class, 12.5 (CI -0.37, 25.5) for the Nonfiction group, and 18 (CI 4.41, 31.7) for the Fiction group (p-values 0.045, 0.057, and 0.009, respectively). Self-reporting of more frequent engagement during medical school correlated with lower scores, with values reaching significance in the Fiction group, but the low n=2 indicates that this finding is not an accurate reflection of the effects of engagement in arts in humanities.

Table 5: Association Between Covariates and Empathy Change				
Variables	Baseline Empathy Change Beta (95% CI)	P-value[1]	Reading Empathy Change Beta (95% CI)	P-value[1]
Genre				
Non-Fiction	REF		REF	
Fiction	-1.00 (-3.48, 1.48)	0.42	1.32 (-1.09, 3.73)	0.27
Age, years				
≤25	REF		REF	
>25	-0.96 (-3.43 (1.51)	0.43	0.72 (-1.67, 3.12)	0.54
Gender				
Female	REF		REF	
Male	-2.93 (-5.39, -0.46)	0.021	-1.25 (-3.65, 1.14)	0.29
Undergrad Multiple Fields				
No	REF		REF	
Yes	-0.02 (-4.44, 4.40)	0.99	3.75 (-0.54, 8.05)	0.087
Undergrad Major				
STEM	REF		REF	
Social Sciences	-1.32 (-4.68, 2.04)	0.43	0.36 (-2.91, 3.63)	0.82
Both	0.01 (-5.01, 5.03)	0.99	-3.71 (-8.60, 1.16)	0.13
Engaged in Arts and Humanities Before Medical School				
Never	REF		REF	
Rarely	5.80 (-5.25, 16.8)	0.29	1.66 (-9.09, 12.4)	0.75
Sometimes	6.22 (-4.39, 16.8)	0.24	-1.01 (-11.3, 9.31)	0.84
Often	5.56 (-4.82, 15.9)	0.28	0.63 (-9.47, 10.7)	0.90
Very Often	5.67 (-4.78, 16.1)	0.28	-4.21 (-14.4, 5.96)	0.41
Engaged in Arts and Humanities while Applying to Medical School				
Never				
Rarely	REF		REF	
Sometimes	1.60 (-6.43, 9.64)	0.68	-1.36 (-9.18, 6.44)	0.72
Often	2.14 (-4.86, 9.14)	0.54	-1.39 (-8.20, 5.42)	0.68
Very Often	1.03 (-5.14, 7.19)	0.74	-0.24 (-6.24, 5.74)	0.93
	N/A			
Engaged in Arts and Humanities While in Medical School				
Never	REF		REF	
Rarely	-1.47 (-7.68, 4.74)	0.63	1.07 (-4.97, 7.11)	0.72
Sometimes	-2.12 (-8.45, 4.23)	0.51	1.41 (-4.76, 7.57)	0.64
Often	-1.00 (-8.90, 6.89)	0.80	0.69 (-6.99, 8.37)	0.85
Very Often	0.36 (-10.6, 11.3)	0.94	4.02 (-6.62, 14.7)	0.45
Baseline Scores	-0.51 (-0.76, -0.27)	<0.001	0.10 (-0.13, 0.34)	0.38

[1]Linear regression to ascertain mean differences in the change in scores between time points adjusting for all other variables in the model.

Table 6: Mean Differences in Scores Over Time

Variables	Overall		Non-Fiction		Fiction	
	Beta (95% CI)	P-value[1]	Beta (95% CI)	P-value[2]	Beta (95% CI)	P-value[3]
Genre						
Non-Fiction	REF		N/A		N/A	
Fiction	0.002 (-2.17, 2.18)	0.99				
Age, years						
≤25	REF		REF		REF	
>25	-0.38 (-2.55, 1.78)	0.72	2.23 (-1.21, 5.68)	0.20	-3.49 (-6.32, -0.65)	0.016
Gender						
Female	REF		REF		REF	
Male	-2.75 (-4.92, -0.59)	0.012	-3.80 (-7.05, -0.56)	0.021	-0.57 (-3.58, 2.43)	0.71
Undergrad Multiple Fields						
No	REF		REF		REF	
Yes	-0.17 (-4.02, 3.69)	0.93	-1.12 (-7.58, 5.32)	0.73	2.56 (-2.08, 7.19)	0.28
Undergrad Major						
STEM	REF		REF		REF	
Social Sciences	-0.34 (-3.29, 2.61)	0.82	-0.61, -6.49, 5.25)	0.83	0.60 (-2.87, 4.08)	0.73
Both	-0.11 (-4.51, 4.28)	0.96	-0.11 (-6.60, 6.38)	0.97	-1.29 (-6.87, 4.27)	0.65
Engaged in Arts and Humanities Before Medical School						
Never	REF		REF		REF	
Rarely	5.41 (-4.29, 15.1)	0.27	9.14 (-7.46, 25.8)	0.28	10.6 (-3.12, 24.4)	0.13
Sometimes	8.30 (-0.92, 17.5)	0.078	10.2 (-2.63, 23.1)	0.12	15.3 (2.09, 28.4)	0.023
Often	9.21 (0.23, 18.2)	0.045	12.5 (-0.37, 25.5)	0.057	18.0 (4.41, 31.7)	0.009
Very Often	7.97 (-1.05, 17.0)	0.083	9.24 (-2.07, 20.6)	0.11	20.6 (6.79, 31.4)	0.003
Engaged in Arts and Humanities while Applying to Medical School						
Never	REF		REF		REF	
Rarely	1.56 (-5.49, 8.62)	0.66	-1.76 (-16.6, 13.1)	0.81	-8.35 (-19.9, 3.23)	0.16
Sometimes	-1.67 (7.77, 4.42)	0.59	-3.71 (-12.7, 5.23)	0.41	-11.9 (-23.3, -0.40)	0.043
Often	-1.95 (-7.31, 3.40)	0.47	-4.05 (-12.0, 3.91)	0.31	-10.6 (-20.3, -0.98)	0.031
Very Often	N/A		N/A		N/A	
Engaged in Arts and Humanities While in Medical School						
Never	REF		REF		REF	
Rarely	-2.86 (-8.27, 2.56)	0.30	-4.63 (-11.7, 2.40)	0.19	-2.49 (-7.06, 2.07)	0.28
Sometimes	-3.26 (-8.76, 2.25)	0.25	-5.07 (-11.5, 1.35)	0.12	-2.16 (-7.06, 2.73)	0.38
Often	-3.05 (-9.94, 3.84)	0.38	N/A		-14.9 (-25.6, -4.14)	0.007
Very Often	-1.78 (-11.3, 7.78)	0.71	N/A		-16.5 (-30.3, -2.80)	0.018
Time Points						
Baseline	REF		REF		REF	
Pre-Read	-0.83 (-1.93, 0.26)	0.14	0.08 (-1.57, 1.72)	0.92	-1.76, (-3.17, -0.36)	0.014
Post Read	-2.15 (-3.25, -1.05)	<0.001	-1.95 (-3.61, -0.30)	0.021	-2.35 (-3.76, -0.95)	0.001

[1]Linear Mixed Model to ascertain mean differences in scores over time.

[2]Linear Mixed Model to ascertain mean differences in scores over time in the Nonfiction subgroup.

[3]Linear Mixed Model to ascertain mean differences in scores over time in the Fiction subgroup.

Discussion

Empathy scores decreased over time for the Overall class, Fiction, and Nonfiction groups. This is consistent with prior studies that demonstrate overall empathy decreases during medical school but does not support previous research on interventions using literary studies found to increase or retain empathy in medical students. However, there are many challenges in assessing the effectiveness of these interventions. Studies that have found positive results in empathy-building/retention have often utilized voluntary participation and more subjective empathy scoring systems based on personal evaluation of one's ability in perspective-taking or empathising.[11, 13, 18, 19] This project has aimed to avoid potential selection and confirmation bias by incorporating the reading intervention as a required aspect of medical curriculum and using a more objective empathy assessment tool. Therefore, our negative findings may be a reflection of the need for more accurate study design and assessment criteria in this field.

We found our results to be consistent with other studies in terms of gender-based empathy trends that self-identified males scored lower than females across all time points. Further subgroup analysis of variables by gender may be helpful in identifying factors that contribute to this difference.

The other significant finding was that increase Engagement in the Arts and Humanities prior to medical school corresponded with higher empathy scores over time. These values hit significance for individuals in the Overall Class who selected Engagement as Often, and those that selected Sometimes, Often, and Very Often in the Fiction group, though the reference value in the Fiction group had only n=1 and is therefore not an accurate representation. In contrast, higher Engagement in Arts and Humanities while applying to and during medical school correlated with lower empathy scores over time. This correlation was not found to be significant for the Overall Class nor for the Nonfiction Group but did reach significance in the Fiction group. This is again not taken as an accurate finding due to low representation in both high-end and low-end (reference) values.

Limitations

It is also important to note that the findings are limited by assessments at only three time points with a single reading intervention that consisted of a 30-page assignment over a period of 11 weeks. Additionally, the intervention was set up for the participants' start of the second year of medical school, which is more than two-thirds of the way through an 18-month preclerkship curriculum. The strategic placement of this intervention was aimed to counter the stressors from the first board exam taken at the end of second year. However, it has been observed that students are already preoccupied with board exam preparations as early as halfway through the first year of medical school, and it is likely that the decreases in empathy scores demonstrated in this study over-represents this stressful process and under-represents the effect of the intervention. Moving the intervention to portions of the curriculum with fewer standardized exams (such as a fourth-year elective/selective) may overcome this challenge, and more longitudinal assessments may demonstrate a more accurate reflection of empathy changes over time during medical school.

Differences, or lack thereof, in outcomes between the Nonfiction and Fiction groups were also difficult to confirm due to absence of data on the actual reading selections that participants chose. Despite randomization into each genre, students may have consulted peers on reading recommendations and chosen selections outside of their randomized group. The degree to which participants completed the 30-page reading assignment is uncertain as well, as no content material was assessed to determine thoroughness or reading completion.

Observations from the Debriefing Session

Although participants engaged in a discussion session of the literary material as part of the curriculum, the Post-reading empathy assessment was conducted immediately prior to the session in order to avoid confounding factors of benefit from the discussion rather than the text itself. Feedback from the session was generally positive, and debriefing and active analysis of the reading material likely contribute more to empathy skills than reading alone. However, participants did struggle more with Fiction than Nonfiction. Students gravitated to selections

with more obvious medical themes, but still found the Fiction narratives less relatable to clinical scenarios despite diagnostic possibilities and interpretations of disease pathology from the text.

Phenomenology

Nonfiction and Fiction selections did have potential to be paired from each genre, which may be useful for future interactions of this activity. In this way, students can relate to scenarios from Nonfiction, apply skills for bias awareness, perspective-taking, and empathic concern to the Nonfiction piece, and then translate and further develop these skills in a Fiction piece with a similar medical theme. These are the skills in medical humanities that have been identified with empathy building. Students may feel challenged more with Fiction (even pieces with medical themes) due to its ambiguous language and abstract concepts compared to Nonfiction, despite both genres having the ability to be "writerly" texts.[13] It is this exposure to unfamiliar representation of familiar (or unfamiliar) situations that can help students develop the previously mentioned skills. Used together, Nonfiction and Fiction can be used to establish a dialogue between the familiar and unfamiliar, a process Arno Kumagai and Delese Wear call "estrangement".[20] Whereas their definition focuses on developing new perspectives of familiar encounters (a skill necessary for gaining value from Narrative Nonfiction), it is equally important to establish comfortable with interpreting unfamiliar circumstances (a skill necessary for appreciating Literary Fiction).

Conclusions

Interventions utilizing literary studies have previously been shown to build or retain empathy in medical students, but this was finding was not supported in this study that incorporated a single reading intervention into medical curriculum. More longitudinal studies and understanding of the skill development provided by the humanities needs to be conducted to better understand how this field should be best-utilized for empathy retention in medical curricula. Integration of health humanities should start at the pre-medical level and continue with training.

Additionally, given the positive feedback from the discussion sessions, it may be more important to pair reading with active analysis in order to help students develop the necessary skills for empathy-building. Careful selection and paring of Fiction and Nonfiction reading options can perhaps maximize the potential benefit from each genre.

References

1. Neumann M, Edelhäuser F, Tauschel D, Fischer MR, Wirtz M, Woopen C, Haramati A, Scheffer C. Empathy Decline and Its Reasons: A Systematic Review of Studies with Medical Students and Residents. *Academic Medicine*. 2011;86(8):996-1009 10.97/ACM.0b013e318221e615.

2. Jones, AH. Literature and medicine: an evolving canon. *The Lancet.* 1996; 348(9038):1360 - 1362.

3. Foster W, Freeman E. Poetry in general practice education: Perceptions of learners. *Fam Pract*. 2008;25(4):294-303.

4. Shapiro J, Rucker L. Can poetry make better doctors? Teaching the humanities and arts to medical students and residents at the university of California, Irvine, College of Medicine. *Acad Med*. 2003;78(10):953-957.

5. Garden R. Distance learning: Empathy and culture in Junot Diaz's "wildwood". *J Med Humanit*. 2013;34(4):439-450.

6. Rosenthal S, Howard B, Schlussel YR, et al. Humanism at heart: Preserving empathy in third-year medical students. *Acad Med*. 2011;86(3):350-358.

7. Wershof Schwartz A, Abramson JS, Wojnowich I, Accordino R, Ronan EJ, Rifkin MR. Evaluating the impact of the humanities in medical education. *Mt Sinai J Med*. 2009;76(4):372-380.

8. Bleakley A. Stories as data, data as stories: Making sense of narrative inquiry in clinical education. *Med Educ*. 2005;39(5):534-540.

9. Bleakley A, Marshall R. Can the science of communication inform the art of the medical humanities? *Med Educ*. 2013;47(2):126-133.

10. Holmgren L, Fuks A, Boudreau D, Sparks T, Kreiswirth M. Terminology and praxis: Clarifying the scope of narrative in medicine. *Lit Med*. 2011;29(2):246-273.

11. Chen I, Forbes C. Reflective writing and its impact on empathy in medical education: Systematic review. *J Educ Eval Health Prof*. 2014;11:20.

12. Kaptein AA, Koopman JJ, Weinman JA, Gosselink MJ. 'Why, why did you have me treated?': The psychotic experience in a literary narrative. *Med Humanit*. 2011;37(2):123-126.

13. Kidd DC, Castano E. Reading literary fiction improves theory of mind. *Science*. 2013;342(6156):377-380.

14. Davis, M. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*. 1983;44(1), 113-126.

15. Bellini LM, Shea JA. Mood change and empathy decline persist during three years of internal medicine training. *Academic Medicine*. 2005;80(2):164.

16. Colliver JA, Conlee MJ, Verhulst SJ, Dorsey JK. Reports of the decline of empathy during medical education are greatly exaggerated: A reexamination of the research. *Academic Medicine*. 2010;85(4):588-93.

17. Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I. The “reading the mind in the eyes” test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*. 2001;42(2):241-251.

18. Spreng RN, McKinnon MC, Mar RA, Levine B. The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment*. 2009;91(1):62-71. doi:10.1080/00223890802484381.

19. Graham J, Benson LM, Swanson J, Potyk D, Darahta K, Roberts K. Medical Humanities Coursework Is Associated with Greater Measured Empathy in Medical Students. *The American Journal of Medicine.* 2016; 129(12):1334-1337.

20. Kumagai AK, and Wear D. “Making Strange”: A Role for the Humanities in Medical Education. *Academic Medicine*. 2014;89:973-977. doi: 10.1097IACM.0000000000000269.